

After the Crashes Are Counted...

Sponsored by:

Task Force on Development of Highway Safety
Manual (ANB25T)

Committee on Statistical Methodology and
Computer Software in Transportation Research
(A5O11)

Workshop Plan

- Four presentations
 - "Why are we here?"
 - Forrest Council, UNC Highway Safety Research Center and BMI-SG
 - "Limitations to modeling and analysis techniques"
 - Simon Washington, University of Arizona
 - "Multivariate Modeling in Road Safety"
 - Ezra Hauer
 - "Bayes' methods,"
 - Dominique Lord, Texas Transportation Institute
- Questions/discussion after each
- Discussion after all four

- Main three talks to concentrate on *statistical methods*
- My initial presentation on
 - Why are we here?
 - Safety *data* issues to consider, regardless of methodology used

Why Are We Here?

- HSM is compilation of knowledge and tools to assist profession in making highway safety decisions
- Knowledge and tools based on research
- For knowledge and tools to be good, research must be good
- Thus, HSM is trying to build in “quality control”

HSM Research Quality Control

- Detailed discussions of primary MOEs
 - Crashes and crash severity
- Research Subcommittee discussion of methods, but none prescribed
- Scientific Review Panel to review study and make recommendation to Task Force
- Input to research funders concerning needed research and review of methods
- Assistance/guidance to researchers
 - This workshop

But Why Is This Necessary?

- Safety research for decades
- Multiple predictive models developed
- Crash Reduction Factors develop and in use by most states
- So why this emphasis?

We Need to Do Better...

- IHSDM efforts

- Methodology: Develop predictive models which would be combined with Accident Modification Factors (AMFs) to predict safety for any situation
- Goal: Base AMFs on well-done before/after studies in the literature (e.g., lane widening, shoulder paving, curve flattening)
- Expert panel found very, very few well-done BA studies. Had to base many AMF estimates on cross-sectional studies (models) where nothing really had been “changed”.
- Base models also had to be redone using latest modeling techniques

We Need To Do Better (cont)

- Recent review of key literature on Red-Light Camera effects
- 16 Before/After and cross-sectional studies reviewed (almost all done after 1995)
- 14 were felt to have significant methodological weaknesses that made results questionable

We Need To Do Better (cont)

- Hauer's WebPage (<http://www.roadsafetyresearch.com>)
- Lane width research since 1985
 - 17 studies
 - About half were felt to have significant methodology problems
- Shoulder width, paving research since 1985
 - 13 or more studies
 - Again, close to half had problems
- Horizontal curvature
 - 10 or more studies since 1985, plus earlier studies
 - Still do not know whether curve is a “point risk” (risk at entry and exit points) or increases risk along full length

We Need To Do Better (cont)

- “Fishing for Safety Information in the Murky Waters of Research Reports” (2002)
 - Spotting Regression to the Mean
 - Use/misuse of significance testing
 - Problems in functional forms chosen for variables in models
- The good news – studies since mid-90s have fewer problems.

Our Hope

- When similar reviews are done in 2013, the same problems will not be found
- That new research will be significantly better
 - HSM
 - IHSDM
 - NCHRP and FHWA
 - CRF work under F-SHRP
 - State- and privately-funded research

Major Data Issues, Regardless of Methodology

- AADT is strongest predictor of crashes
- Most models are based short homogeneous sections
 - Average homogeneous section length in HSIS states – 0.11 to 1.0 miles
 - Shorter on non-Interstate roadways where problems are
- We assume that AADT is accurate for each record (segment) we analyze, and for each year

Major Data Issues – AADT (cont)

- Traffic counts
 - Permanent count stations counting full time all year, but less than 100 in most states
 - Short-term counts (e.g., 48-hour) done on other parts of roadway system annually
 - Short-term counts usually cover a state system in two-five years
- Average count characteristics in HSIS states
 - One annual count per 1-9 miles, depending on state
 - One annual count per 1-28 segments (records)
- Conclusion – Even ADT is an estimate on our research records

Major Data Issues – Crash Locations

- We link crash counts with roadway sections to develop models
- We assume that the crashes on each section record have the inventory and volume attributes on that record
- But we know there are crash location problems
- The shorter the section record analyzed, the greater the chance of crashes being on the wrong records
- Remember that average homogeneous section length in HSIS states – 0.11 to 1.0 miles, with shorter sections on non-Interstate roadways

Major Data Issues – Crash Locations (cont)

- Solution?
- Consider combining adjacent sections into longer analysis records
- But you must then choose which variables are unimportant enough to be combined over
- Choose carefully

Data Improvements to Fight For...

- More complete data on
 - Injury level
 - Sequence of events
 - Exposure by vehicle/person type
 - Horizontal/vertical alignment inventory
 - Roadside and intersection inventories
 - Construction zone crash and inventory info.
 - Urban inventory information

Conclusions

- We need to do better in our analyses, both before/after and modeling efforts
 - Regardless of how strong the method, researcher must
 - Define what data is critical
 - Know the databases that might include them
 - Understand the data issues that can cloud your findings
- “Marry your data (or at least live with it for a while), rather than taking it out for one blind date!”